Assessment: Monitoring & Evaluation in a Stabilisation Context

Joint Doctrine Note x/10

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
15 SEP 2010	**Summary**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Assessment: Monitoring & Evaluation in a Stabilisation Context	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Development Concepts and Doctrine Centre (DCDC), Shrivenham, UK**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **USJFCOM J9 Joint Concept Development and Evaluation, Suffolk, VA**	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This doctrine note amplifies the guidance given in JDP 3-00 on assessment drawing on emerging thinking and practice from different communities on how to monitor and evaluate within stabilisation environments as part of a comprehensive approach. Its purpose is to provide practitioners within a stabilisation context with guidance that will help them structure their thinking, plan and conduct assessments within a stabilisation environment. The document presents key issues that those commissioning, planning and conducting assessments need to consider and provides a range of commonly used options from which they may choose.

15. SUBJECT TERMS
multinational. experimentation, MNE, monitoring and evaluation, comprehensive approach, OECD-DAC, theory of change, theory of conflict

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**51**	

Preface

1. **Purpose**. This doctrine note amplifies the guidance given in JDP 3-00 on assessment drawing on emerging thinking and practice from different communities on how to monitor and evaluate within stabilisation environments as part of a comprehensive approach. Its purpose is to provide practitioners within a stabilisation context with guidance that will help them structure their thinking, plan and conduct assessments within a stabilisation environment. The document presents key issues that those commissioning, planning and conducting assessments need to consider and provides a range of commonly used options from which they may choose.

2. **Readership**. The JDN is written specifically for a UK governmental audience both military and civilian but is intended to be relevant to a range of partners across the international community. It is for use by key decision makers and staff who are engaged directly, or indirectly, with the monitoring and evaluation of stabilisation operations including those deploying as part of the wider HMG engagement from the FCO, MOD, DFID and the UK Civilian Stabilisation Group. The note assumes a reasonable level of knowledge of stabilisation issues but limited expertise in monitoring and evaluation.

3. **Structure**. Section 1 situates the Note within a stabilisation context and outlines the current doctrinal positions and perspectives of key partners. Section 2 introduces the key concepts that shape an effective Monitoring & Evaluation (M&E) regime and outlines the elements that make up a M&E framework. While Section 2 inevitably presents an idealised perspective Section 3 introduces the very real challenges that any M&E approach will face within a stabilisation environment. Section 4 takes the idealised concepts in Section 2 and describes how they can be applied to the challenges outlined in the previous section setting out a framework for the design and implementation of a M&E regime in a stabilisation environment. Section 5 adds further details on specific tools and techniques.

4. **Linkages**. This JDN builds on the guidance given in JDP 3-00 “Campaign Execution”, JDP 3-40 “Security and Stabilisation: the Military Contribution”, JDP 5-00 “Campaign Planning” and Army Doctrine Note 09/07 *“Measuring the Effectiveness of Operations”*. It also has close links with JDN #/10 “Security Transitions”. It draws from a number of civilian publications including the OECD-DAC Guidance on Evaluating Conflict Prevention and Peacebuilding Activities. A Bibliography can be found at Annex #.

CONTENTS

SECTION 1 – INTRODUCTION

The real world is one of constantly shifting environments and constant adaptation to those shifts. This is particularly true in the setting of deep rooted conflict and violence. The most realistic, as in the most realpolitik, thing we could do in peacebuilding would be to create processes with peripheral vision, capable of maintaining purpose while constantly adapting to the difficult and shifting tides and sands they must face to survive. (Lederach 2005)

Stabilisation

101. JDP 3-40 *Security and Stabilisation: the Military Contribution"* describes stabilisation as the process that supports states which are entering, enduring or emerging from conflict. It aims to: prevent or reduce violence; protect the population and key infrastructure; promote political processes and governance structures. This should enable a political settlement that institutionalises non-violent contests for power; and prepares for sustainable social and economic development. It puts forward three central ideas: that stabilisation must be approached "**comprehensively**" across security, governance and development domains. Secondly that the central conflict relationship is that between the host government, competing elites and the wider **population**. Thirdly that the national strategic aim of stabilisation interventions should be to foster the development of a **political settlement**, amenable to UK interests, between this triumvirate of actors. The Stabilisation Unit perspective is similar stating that stabilisation is the process of establishing peace and security in countries affected by conflict and instability. It is the promotion of a peaceful political settlement to produce a legitimate indigenous government, which can better serve its people.

102. Stabilisation will invariably sit within a broader state-building and peace-building context that will seek to address the causes and effects of conflict and fragility and build conflict resolution mechanisms; support inclusive political settlements and processes; develop core state functions and respond to public expectations.

Significance and Purpose of Monitoring & Evaluation

103. As a mechanism for **institutional learning** M&E helps us to systematically improve our understanding, test and refine our assumptions, adapt our strategies and improve our interventions. The empirical evidence and interpretations it provides assists us in drawing convincing conclusions about what works, what does not, how well and why. This helps decision-makers to set priorities, allocate resources and to decide on what efforts should be reinforced, retained, replicated elsewhere, altered or abandoned.

104. As a **planning tool** it can also improve the original plan and facilitate a comprehensive approach. Including M&E within the planning process forces planners to clarify their logic, testing and refining their thought processes and reducing ambiguity. Set within a comprehensive approach, a shared M&E system offers key stakeholders a platform for **dialogue** through which mutual understanding can be developed and differences explored. This builds links between key actors and reinforces effective and coherent action on the ground.

105. M&E can be used as an **accountability tool** to demonstrate the effective use of resources. As such it can encourage transparency and model good governance practices.

As a tool for **advocacy** it can help build a case for continued support or for a change in direction. The data can also be used to **create effects** particularly as a source for strategic communication and public relations.

106. M&E is not an end in itself, but a means to inform decisions. It should be regarded as an iterative process through which our understanding deepens and our interventions improve. The alternative is to rely on preconceived ideas, guesswork or anecdotes. This leads to strategies that lack consistency and that fail to learn; assumptions remain untested, mistakes are repeated and flawed logic endures.

The Principles of M&E

107. To achieve these aims M&E needs to accord with four principles:

a. **Utility**. M&E is not an end in itself, the data and analysis that it produces must have utility; it must meet the purpose for which it has been designed and enable decision-making.

b. **Validity**. M&E must be designed such that the data and analysis portrays what it is intended to portray. The conclusions must be valid.

c. **Achievable**. M&E must be designed such that it can be delivered despite the challenges of the stabilisation environment.

d. **Conflict Sensitivity**. M&E is an intervention in itself. It must be designed in such a way that it does not exacerbate conflict and where possible supports stabilisations efforts.

Current Perspectives

108. Stabilisation doctrine stresses the importance of a comprehensive approach. It is only by drawing together the expertise of a wide range of actors that it is possible to address the breadth of the problem. Collaborative analysis provides an essential shared foundation for such an approach and shared M&E assessments maintain and build on this foundation. It is important therefore to understand how these multinational and interagency partners understand M&E.

Current UK Military Doctrine

109. Current doctrine describes what assessment is from a military perspective but gives limited guidance on how to conduct it, particularly within a stabilisation environment. Although discussed in JDP 5-00 and in JDP 3-40 assessment is primarily covered in JDP 3-00 *"Execution"*. This defines assessment as:

> *"The evaluation of progress, based on levels of subjective and objective measurement in order to inform decision-making".*

110. It is described as a means of keeping the situation and campaign progress under continual review in order to support subsequent decisions. Existing doctrine identifies 3 categories of assessment:

a. **Measurement of Activity**. Defined as the "assessment of task performance and achievement of its associated purpose", it focuses on answering whether "we did properly the things we planned to do" and informs decisions on whether activity should be repeated or altered.

b. **Measurement of Effect**. Defined as "the assessment of the realisation of specified effects"; it considers what effects, intended and unintended have been realised focusing on answering whether "we did the right things". It is used to monitor progress and highlight setbacks and supports current and imminent planning decisions.

c. **Campaign Effectiveness Assessment**. Defined as the "evaluation of campaign progress based on levels of subjective and objective measurement in order to inform decision-making". It considers the timely progress of the campaign focusing on whether "the right things, done properly, are getting us where we want to go or need to be within the required timescale".

111. Measurement of activity and measurement of effect inform reviews of progress against current orders whereas campaign effectiveness assessment informs decisions on what to do next. JDP 3-00 stresses the art of assessment and the role of military judgement based on sufficient information. It cautions against it becoming a mechanistic and overly bureaucratic process that absorbs a disproportionate amount of staff time.

NATO

112. AJP 3 *"Allied Joint Operations"* is the main source of NATO doctrine on Assessment. Three types of assessment are considered:

a. **Strategic Assessment**. This answers the question "is the Alliance achieving its strategic objectives?".

b. **Operational Assessment**. This looks out beyond 10 days asking whether the Joint Force is achieving its objectives and whether the adversary or the Joint Force is likely to achieve its objectives first.

c. **Combat Assessment**. This is carried out by Component Commanders (EG Air, Maritime, Land) and explores whether component commanders are doing the right things to achieve the Joint Force Commander's intent and end-state.

113. NATO doctrine also emphasises the importance of lessons learned stating that its purpose is to learn efficiently from experience and to provide validated justifications for amending existing practice in order to improve performance both during the course of an operation and for future operations.

US

114. Assessment is covered in Joint Publication 5-0 *"Joint Operation Planning"* and in Joint Publication 3-07 *"Stability Operations"*. It is described as the continuous monitoring and evaluation of the current situation and the progress of a joint operation towards mission accomplishment. It deliberately compares forecasted outcomes with actual events to determine the overall effectiveness of the force's employment. It is seen as being particularly difficult within a stabilisation environment and JP 3-07 states that a commander may need to dedicate greater effort to assessment than in more straightforward operations. It also encourages collaboration and consultation between military and civilian actors. Both publications stress that assessment should be considered at all stages: *"assessment precedes and guides every activity within the joint operation process and concludes each operation or phase of an operation"*[1]. The publications also warn that commanders and staff should avoid excessive analysis noting that excessive time and energy spent developing elaborate assessment tools and graphs squanders resources better dedicated to other elements of the operational process. Assessment is seen as comprising three activities: monitoring the current situation to collect relevant information; evaluating progress towards attaining end state conditions, achieving objectives and performing tasks; recommending or directing action for improvement.

115. Monitoring is defined as *"the continuous observation of those conditions relevant to the current operation"*. Evaluation is defined as *"the use of criteria to judge progress towards desired conditions and determining why the current degree of progress exists"*. JP 5-00 emphasises the importance of rationalising the problem constructing the guiding logic that will unravel it and identifying the mechanism that will achieve success. In terms of criteria US doctrine refers to both measures of effectiveness (MOE) and measures of performance (MOP).

a. **MOEs**. These are used to assess changes in system behaviour, capability or the operational environment that measure the attainment of an end-state or objective; changes might be positive or negative. MOEs help answer the question "are we doing the right things?"

b. **MOPs**. These are used to assess friendly actions and are tied to measuring task accomplishment. They answer the questions "was the action taken?" and "were the tasks completed to standard?" and help to determine whether "we are doing things right".

116. Both MOEs and MOPs use indicators defined as *"an item of information that provides insights into a MOE or MOP"*. A single indicator can inform multiple MOEs and MOPs however the doctrine suggests that MOEs should draw on impact indicators.

Current UK Stabilisation Unit Guidance

117. Place Holder for SU Guidance

---

[1] JP 5-00 Chapter III, Section D Paragraph 18c.

Other Civilian Perspectives

118. The majority of civilian organisations such as donors, UN agencies and academic institutions use “assessment” to refer to what the military generally think of as analysis. What the military call assessment civilian organisations usually refer to as M&E. Although different definitions exist the most widely accepted terms are those used by the OECD-DAC.

a. **Monitoring**. A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an intervention with information regarding the use of allocated funds, the extent of progress, the likely achievement of objectives and the obstacles that stand in the way of improved performance

b. **Evaluation**. The process of determining merit, worth or value of an activity, policy or program. It consists in the systematic and objective assessment of an on - going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision making process of both recipients and donors.

119. Civilian organisations also differentiate between formative and summative evaluations. Formative evaluations take place during the course of an intervention and are designed to improve the intervention as it proceeds. Summative evaluations are generally conducted after an evaluation has concluded and examine to what extent the desired impact was achieved and value derived; its aim is to derive insights to inform more effective programming in future interventions. These in turn can be further subdivided into many different types of evaluation such as joint, participatory, theory-based; these are described in the glossary at Annex # and discussed further in Section 4.

Terminology used in this JDN

120. The terminology used in this JDN is largely based on that used by civilians so as to facilitate a comprehensive approach.

121. For the purposes of this JDN, Monitoring refers to the process of tracking changes in the environment. It is the continuous process of gathering and interpreting information to maintain awareness, identify what is happening in the environment and what activities have been conducted, are underway or are planned by a wide range of actors. It is the activity that allows one to describe the current situation. Monitoring helps identify to what extent a plan has been implemented and set benchmark or threshold values have been achieved. It may also track the status of critical assumptions identified within the planning process.

122. Evaluation refers to the interpretation and explanation of why and how observed changes are caused (or not) and to assessments of the quality of interventions. It draws out deductions from monitoring. Evaluations seek to explain the course of the intervention, whether activities have been carried out correctly, whether higher level outcomes and impacts are being achieved and whether the right things are being done to achieve the desired change.

SECTION 2 – ASSESSMENT CONCEPTS & TERMINOLOGY

Assessment Concepts

201. Effective M&E requires an understanding of key concepts. It is crucial to understand and articulate the problem that one is addressing as well as the proposed solution, while recognising the assumptions and logical links that make up this understanding.

Understanding the Problem: Theory of Conflict

202. An up to date conflict or problem analysis is an essential precursor to any evaluation. The problem and the context need to be analysed and the understanding gained captured. A clear expression of the problem and the actors, factors or drivers that are influencing it both positively and negatively is required. Developing this understanding collaboratively with others will provide a richer, more informed, understanding and contribute to effective collaborative action later on.

203. The set of assumptions and hypotheses that make up this understanding form a theory, or possibly theories, of the conflict and offer an explanation of the problem. This theory should express the underlying causes of the problem and those actors and factors that are either driving it or mitigating it as perceived by the planning team at the time. Expressing the theory of conflict helps planners develop the logic that will shape their response and deduce the objectives of the intervention. While steps 1 and 2 of the military estimate process cover this analysis the emphasis tends to be on drawing out the deductions rather than capturing the implicit logic; both need to be done.

204. This theory will however be flawed. Information will be lacking, incomplete or wrong and a number of inaccurate assumptions will have been made. This theory therefore forms an initial baseline understanding and will need to be reviewed, updated and amended as better information becomes available and as understanding deepens. M&E provides the mechanism for doing this.

Understanding the Solution: Change Logic

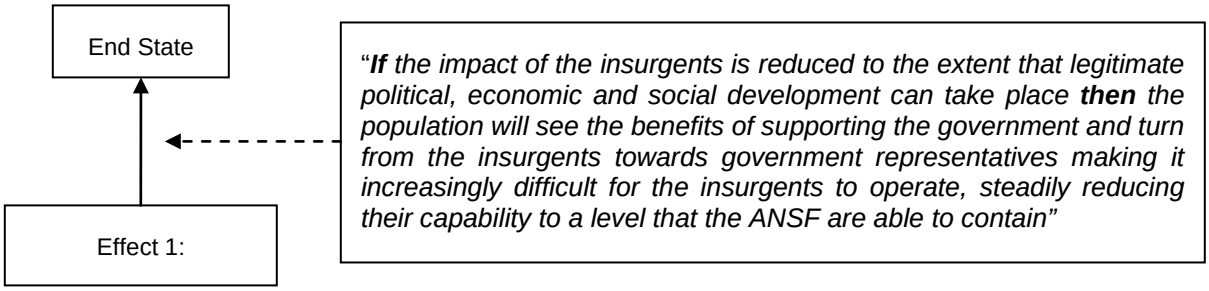
205. The theory of conflict will suggest ways in which the problem might be addressed. These potential solutions contain an implicit "Change Logic", a *"set of beliefs about how change happens*"[2]. This is the logic that underpins the solution and explains how and why the planning team thinks the proposed actions will change the situation. It applies at the macro and the micro level.

a. **Macro Level**. At the macro level it describes generic approaches to conflict; the stabilisation model described in JDP 3-40 is a macro level change logic, a list of others is included at Annex A. A stabilisation campaign is likely to draw on several macro level logics to address different aspects of the overall problem. When combined these form the overall description of the stabilisation campaign and explain why it is thought that the

[2] Church and Rogers

approach selected will create the desired change. For large problems it may aid clarity to articulate the logic of change for different sectors or lines of operation summarising these within an overarching logic.

b. **Micro Level**. At the micro level, it is represented by the connectors that are typically used to join a set of sub-objectives to their higher level objectives. These links summarise the perceived causality between the objectives; they explain why the achievement of sub-objectives A and B is expected to generate a particular change to objective C. These causal chains can be expressed as a hypothesis often in the form "if…then…". The example is taken from a draft of the HQ ISAF 2008 campaign plan.



Understanding the Intervention: Using the Intervention Logic

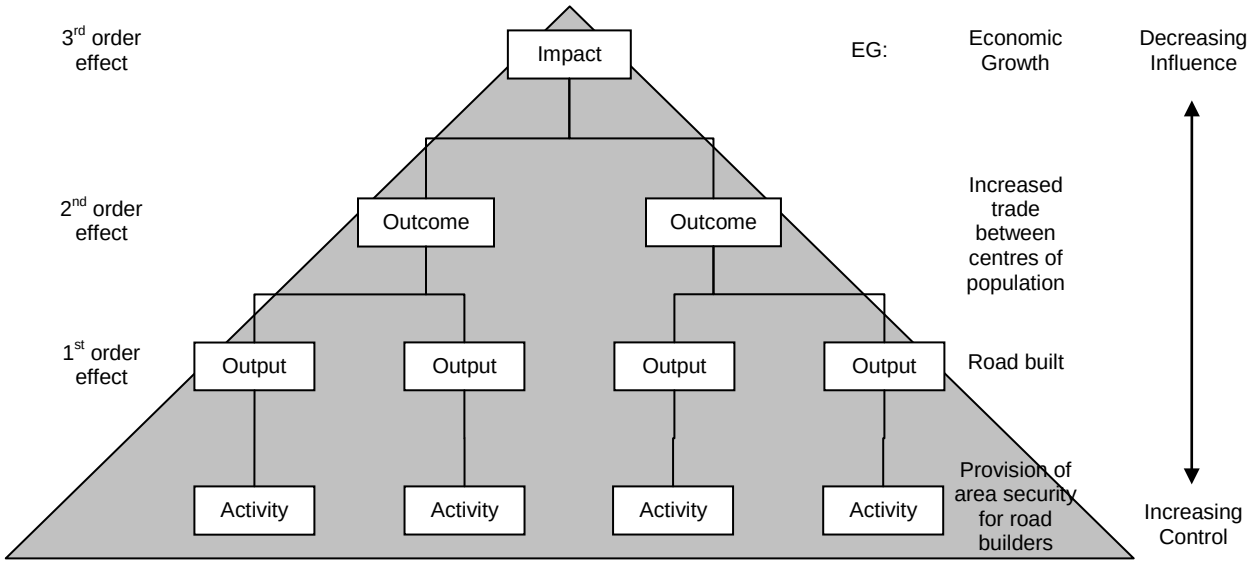
206. Together the Theory of Conflict and the Change Logic form the Intervention Logic. Although critical for planning and M&E, it is rare for the thinking behind them to be captured explicitly. Most hypotheses, logical links, and assumptions tend to remain undocumented. Yet, recording the intervention logic is of use in both planning and assessment:

207. In planning, the discipline of articulating the logic between each step can reveal hidden assumptions as well as highlight inconsistencies or gaps. It forces planners to think through the purpose and logic of the intervention thoroughly. When drawn from a shared theory of conflict it can also help promote coherence enabling a more integrated approach. If left undocumented, different views as to why activities will lead to a desired change will exist, even between members of the same planning team. Misunderstandings between different actors can therefore occur with the original intent being diffused. Writing the logic down and recording the institutional memory is also an invaluable aid to campaign continuity as staff members move on.

208. In assessment, articulating the change logic provides a baseline against which evaluators can make judgements. It provides the collection of hypotheses and assumptions that will need to be tested through M&E to improve the intervention. Thorough records highlight which issues were seen as critical at the time of planning and explain why the intervention is focused where it is. Knowing this, assessment teams can focus their efforts, prioritising those areas which are most significant or where most risk has been taken. Action can then be taken or evidence sought to validate the most important hypotheses, causal links and critical assumptions, thus allowing structured learning and adaptation. Where the intervention logic is not expressed, evaluators and planners should reconstruct it.

Change Logic as a Logical Tree: M&E "Inside the Shape"

209. The Change Logic can be documented in several ways. One of the most common is as a graphic of boxes and connectors forming a logical tree usually making a triangular shape as below. They can also be presented as a table as in logical frameworks[3] or as a narrative.



210. Planning teams are strongly encouraged to articulate each of the links, the causal chains, within this logical tree. Vague or ambiguous causal chains suggest a lack of clarity in planning and make it hard to evaluate and improve that aspect of the intervention. M&E must cover this logical tree – the "shape". It must check progress against the various levels of objectives (the boxes) and test and refine the critical assumptions underpinning the causal chain between objectives (the connectors).

Validating the Theory of Conflict: M&E "Outside the Shape"

211. The logical tree or "shape" is however dependent on the assumptions within the Theory of Conflict. Due to the dynamic nature of the environment and to the inherent flaws in any analysis or planning process, the plan may appear to be progressing adequately while in reality it is failing to gain traction with the problem; outputs may not be achieving the desired outcomes or impact. The theory of conflict may be flawed. It is therefore necessary for M&E to look "outside the shape" as well as "inside". To track changes that are not captured by the logical tree.

Dealing with Complexity: M&E Ourselves

212. Stabilisation environments are both complicated and complex. It is rare therefore to be able to predict events and the overall consequences of actions. Consequently interveners must be adept at sensing and responding to events as they occur. M&E should also be used to assess our own ability to adapt and respond.

---

[3] Usually referred to as "Log Frames"

Bosnia. When General Rose arrived in Bosnia in 1994 as Commander UNPROFOR he found that the headquarters had grown and become overly bureaucratic. One of his first actions was to move the core of the headquarters creating a much leaner and more agile structure better able to respond to events.

213. A stabilisation evaluation plan should therefore include elements that focus on these three areas: the objectives and links within the logical tree, (M&E within the shape), the wider environment (M&E outside the shape) and oneself (M&E assessing ones own adaptability). Regular reviews should be included that examine one's own internal performance, processes and structures to ensure that they remain able to sense and respond rapidly to changes in the environment.

Key Terminology

Efficiency & Effectiveness

214. Evaluations are judged against criteria and it is important to be clear what these are. Efficiency, how well inputs are transformed into outputs, and effectiveness, the extent to which an intervention has achieved its objectives are the most common however, a number of others such as relevance, coverage, coherence, impact and sustainability also exist. These are discussed in Section 4 and described in the glossary.

Cause & Effect

215. In a controlled environment such as a laboratory it is relatively straightforward to identify cause and effect. However, in stabilisation environments the situation will be affected by a wide range of variables that cannot be examined in isolation. As a result it is difficult to state with any certainty what action has caused a certain outcome, let alone how much it has contributed. This becomes more so the higher up the logical tree one moves, particularly as many impacts will take time to appear. It becomes increasingly hard therefore to accredit a certain impact or outcome to a specific input; to say that A caused B which caused C. A more nuanced understanding of causality is required where direct attribution is replaced by a broader assessment of contributions. Relevant terms include causality, attribution, contribution, correlation and generalisation and are described in the glossary.

Quantitative & Qualitative

216. Quantitative approaches attribute a numerical value to data and base their findings and arguments on numerical units. Qualitative approaches use words, images and sounds instead of numbers although they can be grouped and aggregated to provide numerical data. The fact that an approach is expressed in numbers or text does not define it as either quantitative or qualitative; it is important to look behind the expression to see what is actually being measured. If it is an opinion, a belief or a way of thinking it is qualitative data. Neither quantitative or qualitative data or approaches are better or worse, they both have strengths and weaknesses and good assessment frameworks will draw on a mix of both.

Subjective & Objective

217. Objective and Subjective are often used alongside quantitative and qualitative with objective seen as "good" and "subjective" seen as "bad". This is overly simplistic. Objective refers to observable facts, things that exist independently of perception and that are

unaffected by personal perspectives. Subjective brings in perspectives, attitudes and emotions; it refers to how people think about an issue. Ultimately all assessments are subjective as they involve making judgements and the level of subjectivity increases the further an assessment progresses. Deciding what data to collect and what to omit involves judgement and therefore an element of subjectivity exists in the very design of any assessment framework. The meaning that individuals place on any data also draws on their worldview, often unconsciously, and therefore subjectivity increases. Pure "hard" objective data unaffected by perceptions, institutional norms or attitudes is therefore rare. Assessors must develop an understanding of the flaws and weaknesses inherent in any data and ensure that conclusions drawn reflect this; overly simplistic characterisations of data as being objective or subjective can be misleading. The use of quantitative and qualitative and of subjective and objective approaches is discussed further in Section 4.

Elements of an M&E Framework

Baseline

218. A baseline provides a starting point from which a comparison can be made and changes identified. Ideally a baseline should be established before an intervention starts. However this is not always possible. An evaluation planning process should therefore identify what data exists and from that construct a baseline; missing or corrupted data may need to be the focus of early monitoring efforts.

Data

219. Data forms the basic building block of any monitoring process or evaluation. It is the raw information prior to any assessment or analysis. It might comprise measurements, observations, interview records, survey responses, imagery, video etc. It can be qualitative or quantitative. Baseline data refers to data collected before the start of an intervention. Longitudinal data is collected at periodic intervals throughout the life of an intervention.

Information

220. Information is defined[4] as the meaning that an individual associates with data presented in context. Information is processed data. This processing may not be a formal or even conscious process. It occurs when individuals look at data through the lens of their own cultural norms. It is not therefore value free. The photograph of a Royal Marine in Afghanistan carrying a young Afghan girl shown in JDP 3-40 is data. The information that a western observer might deduce from this data is that the Marine is rescuing the girl and bringing her to a place of safety. The information that an adversary might deduce is that a foreign soldier is carrying a young girl away from her family.

Variables, Assumptions and Hypotheses

221. A variable is an issue or object that can change over time. In a logical tree, inputs, activities, outputs, outcomes, and impacts all constitute variables and changes to these can be tracked. Assumptions are also variables. Three types of assumptions can be identified:

[4] JDN 4/06 "*Information Management*" dated June 2006

those inside the shape upon which the change logic is based; those outside the shape upon which the theory of conflict is based; and those outside the shape that concern how other actors may respond. Hypotheses express the change logic underpinning plans, normally as 'if…then' statements (see paragraph 205), they are based on assumptions.

222. Identifying variables of interest is a key aspect of any M&E effort and must flow directly from the purpose of the evaluation and the decisions that the evaluation is meant to inform. There will usually be more variables of interest than it is possible to monitor and assess; prioritisation therefore is vital and is discussed in Section 4.

Indicators & Direct Measurements

223. Changes in some variables will be observable and easy to track, in particular at the bottom end of the logical tree; e.g. the number of fields where opium poppies are grown. In such cases, direct measurements can be used to track changes. In other cases, changes to variables may be observable in principle, but not in practice, perhaps when the security situation restricts access. Others, in particular at the higher level of the logical tree, may not be observable at all. The legitimacy of the host nation government or the perceived security could serve as examples of such so called constructs. Indicators are used to capture changes to unobservable variables. These are observable changes that correlate with or influence the variable of interest.

M&E Standards: Targets, Benchmarks, Threshold

224. In tracking and assessing change a key issue is how much change is required. M&E makes use of standards against which change is assessed. Two types are considered:

a. **Process standards** are used to assess the way in which things are being done and can be applied as a 'yardstick' when examining efficiency; they tend to be based on "best practice" from similar sectors or situations. In fields like financial accounting, health care etc. such quality standards are readily available. In other cases, planners and evaluators may have to draw on their experience and contextual knowledge to define what standards can be considered as adequate and acceptable for particular processes in the given situation.

b. **Achievement standards** are used when the evaluation is results orientated and seeks to identify to what extent the intervention has achieved planned objectives.

Figure 4:1 M&E Standards

225. In both cases targets, benchmarks and thresholds can be used:

a. **Targets** identify the level of change required in a variable in the context of a specific intervention. There are likely to be several targets associated with any variable that together map out the level of change expected over the course of an intervention.

b. A **benchmark** is a reference point against which performance or achievements can be assessed. Benchmarks are usually drawn from achievements in similar situations or by similar organisations; they may be captured in international standards or guidelines. They can be helpful in identifying what is practically achievable or acceptable within a certain situation and can help identify what "good enough" might be.

c. A **threshold** describes a target that, when reached, triggers a reaction. A variable reaching a threshold might act as a warning prompting a planned reaction, enable activities in other areas to start or mark the beginning of the next phase of a plan. It can also mark the level at which any change may become statistically significant.

Indicator	Target	Benchmark	Threshold
Number of people with safe and equitable access to sufficient water	10l per day per person available from a source no more than 1000 m away by 30 Sep 2010	15 l per day per person available from a source no more than 500 m away	Flow from well needs to be 1 m^3/min before building of irrigation system starts.

Line of Inquiry

226. Lines of inquiry provide the general themes that guide an evaluation to ensure that it meets its purpose. They are analogous to the lines of operation in campaign planning.

Means of Verification

227. Means of Verification refer to the ways in which the data will be collected. Wherever possible it is best to use multiple measures and triangulate in order to add rigour and enhance accuracy. Means can include reviewing existing documentation and records, conducting interviews, polls or surveys and observation. These are discussed in Section 5.

Means of Analysis

228. Means of analysis refer to the way in which the data will be analysed. There are two broad approaches to analysis: deductive and inductive. Deductive identifies a series of hypotheses in advance and then seeks to prove or disprove these hypotheses. It tends to be more focused than inductive analysis and the problem set bounded however it runs the risk of missing key insights that were not identified in advance. Inductive analysis is a more open-ended approach that iteratively uncovers issues and themes leading to a deeper understanding of the subject matter. Effective evaluations will normally combine aspects of both deductive and inductive approaches with deductive approaches being focused on key assumptions or causal links. The choice within this then comes down to the balance between quantitative and qualitative data analysis. This is discussed further in Section 5.

SECTION 3 – M&E IN STABILISATION OPERATIONS

The Challenges

301. Evaluating progress in fragile and conflict affected states is difficult; attempting this within a stabilisation environment which is both complicated and complex poses additional challenges.

Complicated and Complex. An aircraft electronic system can be thought of as complicated. While an ordinary observer may not understand how it works this is due to lack of knowledge. Given time and the right handbooks an observer could navigate their way through the complicated electronic systems and sub-systems and predict what effect pressing any particular button would have. An aircraft full of passengers however is complex. Even if it were possible to analyse the character and behaviour of each passenger and to observe every event and interaction that preceded them boarding the plane it would still not be possible to predict their response to a specific event – although with hindsight it might be possible to understand why they reacted in a certain way. Complicated situations are difficult but knowable (given sufficient time and resources); complex situations however are, by definition, unknowable and therefore unpredictable.

302. The key to complicated environments is that given sufficient knowledge they can be resolved and responses to interventions predicted with increasing confidence. Increased knowledge can lead to more accurate assumptions, more effective change logic and ultimately a more effective intervention. Evaluations in complicated environments therefore need to focus on exploring these assumptions and increasing the level of knowledge and understanding of the environment and its dynamics. M&E within and without the shape can help increase this knowledge and develop this understanding.

303. Complex environments however are unpredictable; no matter how much information one gathers responses cannot be predicted with any degree of confidence. In these environments interveners need to be able to identify and respond to patterns quickly; they need to sense what is happening and then respond. Evaluations in complex environments therefore should focus not on the environment but on the intervener. M&E in a stabilisation environment, being both complicated and complex, must address all three areas: within, without the shape and the intervener.

304. The types of challenges created by this complicated and complex stabilisation environment can be considered in terms of security, staffing, organisational and cultural and resourcing challenges. Together these create significant data collection and analysis challenges.

Security Challenges

305. Security challenges will invariably hamper access. Certain areas may prove difficult to gain information from and in many cases few means of verification will be feasible. Even if an evaluation team is able to access an area temporarily it is unlikely that locals will be willing to participate meaningfully and responses are likely to be guarded. Where views are difficult to collect or where security hampers collection it tends to be those with a particular interest that are willing to express them leading to an unrepresntative view. Gaining balanced information across age groups, gender, religious or tribal affiliations may prove virtually impossible. Care should therefore be taken in making generalisations about such data; sample sizes are likely

to be small and differences may not be statistically significant. Poor security can also be used as an excuse by policy makers or local authorities to restrict access to an area or section of society in order to distort findings and cover up poor performance or misdeeds.

Staff, Organisational and Cultural Challenges

306. The turnover of staff within stabilisation environments can be high leading to a lack of corporate memory. Intervention logic is frequently lost and relationships with other organisations interrupted. Such environments can also be characterised by a higher than usual proportion of inexperienced staff.

307. Of the multitude of actors present in such situations many will have different perceptions of the problem and of the solution. While this can prove constructive adding richness and depth, differing interests (not all of which will be overtly stated) will also bring to the fore inter-organisational politics. Competition for funding, a desire for individual and organisational acclaim, concerns over reputation and institutional practice can all combine such that information is seen as a commodity and a source of power; information sharing is then limited and the scope for open and genuine dialogue reduced.

308. This can be exacerbated by a culture of secrecy within organisations. Many organisations may have genuine concerns over maintaining the confidentiality of their sources and ultimately the protection of their staff. The military also have rules regarding the sharing of information on grounds of operational security and will use security classifications to control distribution. While these procedures are required, care must be taken to prevent the over-classification of information; the risks of sharing must be balanced against the risks of not sharing.

Kabul 2008. In HQ ISAF in 2008 information was shared between UNAMA and the headquarters. On one occasion UNDSS had produced some data which was shared with the headquarters; this was then collated and turned into a powerpoint presentation which was classified by the ISAF desk officer. As a result UNDSS were no longer able to see the information that they had provided. This was not an unusual occurrence and undermined relations between the two organisations.

309. In many cultures the whole concept of evaluation or of certain evaluation approaches may be alien being seen as a foreign imposition to be complied with only reluctantly. The idea of someone judging performance can also be seen as threatening.

While the concept of polling and the statistical basis behind it is widely accepted in the West this is not the case elsewhere. In Afghanistan the suggestion that a statistically valid poll from a sample population could be used to draw conclusions about the broader population appears bizarre to many Afghan officials; persuading them to make policy decisions based on such evidence is therefore difficult.

Resourcing Challenges

310. There is an inevitable desire within a stabilisation environment to get on with the task of stabilisation; windows of opportunity may be limited and must be exploited. Investing analysis and planning effort in monitoring and evaluation can be seen as an overhead preventing activity, particularly if those involved in gathering the data have not been involved in the design of the evaluation and if there is little in the way of feedback. Consequently evaluations are often seen as an extractive process that places demands on those at the front end but

which provides them with little in the way of useful insights. As a result data collection can be haphazard, imprecise and of poor quality; inclusion and feedback can mitigate this.

Data & Analysis Challenges

311. As a result reliable data can be hard to access. Baseline data is often incomplete or out of date. Inconsistent engagement by the international community may mean that information is available for certain areas but not others or for only certain periods of time. Records may have been destroyed in the conflict or deliberately falsified to strengthen or undermine a faction's position. Existing collection or reporting mechanisms may have broken down. Such information as does exist may no longer be accurate due to displacement with the presence of IDPs or refugees distorting pre-crisis data. In many situations there may be no agreement over such mundane items as place names leading to ambiguity over where data refers to. In such environments data is also political and will be used and abused to paint a picture to support or undermine different parties' agendas. It is also an asset that can be used as a bargaining chip; open and unconditional sharing may therefore be a rarity.

312. Feedback is also important as participants can suffer from questionnaire fatigue as another group with another questionnaire comes to ask similar questions but do very little. Evaluators should explore which other organisations may be conducting surveys and interviews and wherever possible merge data collection methods or make use of each others' data.

313. The easiest data to acquire therefore is that which is reported upwards from within the organisation using its own resources. The pressures of internal staffing procedures, briefings and the organisational "battle rhythm" exacerbates this encouraging staff to look inwards rather then externally to what is happening in the environment. Internal data also tends to be orientated to the accomplishment of the organisation's tasks and invariably reflects their worldview. The danger is that such reporting therefore simply confirms our own cultural bias.

Conflict Sensitivity

314. An evaluation is in itself an intervention and will affect the environment that is being evaluated. A key principle is that the evaluation should not exacerbate the situation, increase tensions or reinforce the conflict drivers. Evaluators also have a responsibility to protect those that they are drawing information from.

Conflict Dynamics

315. Evaluators must understand the conflict dynamics and consider how their evaluation might affect them. The results of any evaluation can be abused by those in positions of power to further their own ends. Evaluators should consider the timing of interviews or polls and the time, place and manner in which results will be announced. Political events, religious festivals or anniversaries can act as triggers and evaluators will need to factor these in to their evaluation plan. There is also a danger that the attention of external agencies and the nature of the questions posed may raise unrealistic expectations. Managing expectations must be a significant element of any evaluation and evaluators should explain the purpose of the questions and how the data will be used.

Protection of People

316. Evaluators must also consider the security of those that they engage with. In some environments simply being seen talking to an outsider may place an individual in danger. Even if participants' details are hidden maintaining confidentiality is important. It may be relatively straightforward for people, particularly in local settings, to work out who said what, especially if those doing so demand a low burden of proof before acting. Questions may also make individuals recall painful or distressing events that they may not wish to be reminded of or that may reignite psychological trauma; it should always be possible for individuals to refuse to participate.

Contextual Sensitivity

317. Finally evaluation staff must be sensitive to the cultural context. In many societies certain topics may be considered taboo and should not be asked or may need to be asked in a particular way. In some situations the choice of language can be a political issue and will betray a certain perspective, similarly certain words, place names or titles can carry political weight suggesting that those asking come from a particular political perspective. Responses may be tempered to reflect the perceived position of those asking or defining the questions undermining the validity of the findings. The nature of the person asking the questions or collecting the data will also affect the responses. Individuals may respond differently to a person in uniform than to a civilian, to an international worker than a local. Age and gender may also be a factor.

The Peer Ethnographic method[5] was developed to enable research by an NGO into sexual and reproductive health issues in Zambia. The NGO recognised that what people said about aspects of their social life depended on the level of trust they had with the person conducting the interviews. As a result they identified individuals within the peer groups that they were interested in who already had a high level of trust. These peer reviewers were then trained in the interview and data capture techniques. This approach produced high quality qualitative data.

[5] Hawkins and Price (2001) The Method

SECTION 4 – PLANNING AND CONDUCTING M&E

401. Section 4 takes the concepts outlined in Section 2 and describes how they can be applied to overcome the challenges of the stabilisation environment described in Section 3. It describes firstly how to design an M&E framework and secondly how to manage and exploit an ongoing framework. More detailed discussion of tools and techniques is reserved for Section 5.

Planning M&E

M&E as an Integral Part of Planning

402. M&E is an essential element of planning and informs re-planning. It must be considered from the outset and throughout any planning process if it is to fully deliver all the benefits outlined in Section 1. The tendency to defer consideration of M&E must be resisted. M&E is conducted for a purpose. It exists to inform a decision or decisions about current or future interventions. To obtain an effective M&E framework planners and M&E experts must enter a dialogue to ensure that it has utility, can be achieved, is valid, and conflict sensitive.

Utility – Audience and Purpose.

- *Confirm who the primary decision-maker is that the M&E effort is to inform.*
- *Clarify the purpose of the M&E effort.*
- *Identify the key questions that the M&E effort will answer to support the primary decision-maker.*

Achievable – Scope and Collaboration.

- *Identify what other M&E efforts are planned or underway clarifying the potential for collaborative evaluations such that duplication of parallel M&E efforts is avoided.*
- *Deciding who will participate in the M&E effort.*
- *Confirm priorities ensuring that M&E efforts are focused.*
- *Determine the scope of the M&E effort.*
- *Establish the timing, duration and frequency of the M&E effort.*
- *Identifying the Lines of Inquiry.*

Achievable – Designing the Framework.

- *Understanding and prioritising the Intervention Logic.*
- *Establishing the baseline.*
- *Identifying and prioritising the variables of interest.*
- *Establishing indicators and standards.*
- *Selecting means of verification and means of analysis.*
- *Forming the evaluation plan.*

Validity & Conflict Sensitivity – Testing the Framework.

- *Confirm ability to access the data.*
- *Confirm the utility of likely responses.*
- *Check for bias or a lack of balance.*

- *Check for conflict sensitivity.*

403. Although shown as a linear process, M&E planning is, like all planning processes, an iterative process. Insights gained when considering the achievability of an approach may for example force designers to reconsider the purpose. Similarly consideration of indicators and means of verification may lead to changes in the scope of the evaluation.

Utility: Audience and Purpose

404. The first stage of ensuring the utility of an evaluation is to clarify the purpose and audience. It must be clear who is going to use the assessment and for what; the type of decisions they will make as a result of it and the key questions that will need to be answered.

Audience

405. Determine who the main user of the assessment will be and what they intend to do with the results. It is likely that there will be many calls for the products of any evaluation and many interested stakeholders. Evaluators should be clear who their primary client is and differentiate them from other interested parties. While there may be many interested readers there should only be one client. Evaluators should resist the temptation to expand the purpose and resolutely ask what the key questions are that the evaluation should answer. Explicitly identifying the client will help focus and prioritise the assessment. It will also allow evaluators to tailor recommendations and present findings in a form best suited to the user.

In Afghanistan in 2008 HQ ISAF operated a Quarterly Campaign Assessment. Although conducted by HQ ISAF and briefed to COMISAF it had been designed to feed JFC Brunssum's Campaign Assessment. As a result it was not providing the information that COMISAF required to inform the decisions that he needed to make. At the same time COMISAF also needed information to brief the US Congress in order to make a case for additional resources. While the same basic data could be used to serve all three needs the requirements of each were very different. As a result there was confusion amongst the staff as to the purpose and the audience for the assessment.

Specifying the Purpose

406. Designers must identify whether the *primary* purpose of the evaluation is for learning, dialogue, accountability, advocacy or to create effects (see paragraphs 103 – 105). Is the aim to learn from current activities to inform the current or future interventions? To provide a forum for dialogue between actors and enable a more integrated approach? To explain and justify the decisions made and the activities conducted? To argue for additional or different resources or a change of strategy? Or is it to generate a particular effect? Each purpose will lead to a different design. A lack of clarity over the purpose will lead to a confused evaluation framework that is unlikely to meet the needs of the client and fail the utility test.

407. Once the primary purpose of the evaluation is clarified evaluators need to focus in on the *specific* purpose of the evaluation. The OECD-DAC has developed a set of evaluation criteria that can guide evaluators in their thinking and help them formulate the key questions around which their evaluation will be focused, others such as ALNAP[6] expanded on these.

[6] The Active Learning Network for Accountability and Performance in Humanitarian Action

Using the OECD-DAC/ALNAP Criteria

408. These can be broken down into four groups. The first two examine the quality of the intervention logic by assessing how it fits into the larger context, looking 'outside and beyond the shape' and by assessing the validity of its causal links, looking 'inside the shape'. The third group examines how well the plan has been implemented, again looking 'inside the shape'. This Note introduces a fourth category that examines how well the organisation can adapt to changing circumstances, thus looking at our own adaptability.

409. **Larger Context**.

a. **Relevance** asks to what extent objectives and activities respond to the issues most critical to the situation. It asks whether the intervention is scratching the right itch and if so for whom. Where the focus is on the local population it is vital to look from their perspective; are the activities addressing their needs as they perceive them or as defined by others? Do they feel that it is relevant to them? An evaluation based on relevance will examine many of the assumptions within a plan, the underlying theory of conflict and the related logic of change. It is focused on whether the objectives are appropriate and whether the activities are consistent with those objectives. It is possible for an intervention to be efficient and effective but not be relevant.

b. **Sustainability** asks to what extent the intervention generates results that will endure after the intervention ends. It can also investigate how adaptable the results are to change. Can they evolve to meet the demands of new phases in the conflict? It asks whether the intervention has generated sufficient popular support; whether the outcomes are politically and socially acceptable, the processes created appropriate and funding streams established. It takes a longer-term view of the activities conducted and their outcomes, questioning how robust any changes may be.

c. **Coherence** examines how the intervention logic fits with our national objectives and with the host nation's and international community's overall stabilisation efforts. It considers the relationship of the intervention under examination with the wider picture. Within a specific functional area it is concerned with whether activities focused on delivery are consistent with and supported by activities at the management and policy levels. More broadly with whether activities in one functional area are coherent with activities in other areas. How coherent, for example, are efforts to build policing capacity with efforts to develop a judicial capacity? It explores forms of coordination and collaboration and asks whether the resources dedicated to them were sufficient and effective. Related ideas are those of **connectedness** and **linkages** both of which explore the coherency of the intervention across different sectors or domains and with longer-term and related problems.

d. **Coverage** refers to the reach of an intervention identifying who is affected by it, who is excluded and the implications of this coverage on conflict dynamics. It can be a useful check to ensure that an intervention is not fuelling underlying tensions and exacerbating conflict drivers thus ensuring that it is conflict sensitive.

e. **Appropriateness** can be considered a subset of relevance and of sustainability in that it explores whether the intervention is tailored to meet local needs or whether it is skewed towards external perceptions of that need or to external needs.

410. **Validity**.

a. **Impact** explores the positive and negative effects of an intervention. It looks both "inside" and "outside the shape" and identifies how the situation, in particular the conflict drivers, has changed. It combines elements of effectiveness and relevance and sheds light on the accuracy of the theory of conflict and the change logic. The timing of an impact evaluation is important as effects can take significant time to emerge.

b. **Effectiveness** is a measure of the extent to which an intervention has achieved its objectives in a timely manner. It examines whether the intervention is achieving its purpose and if the intervention logic is plausible, complete, and accurate. It therefore focuses on objectives and the causal links between them.

411. **Implementation**.

a. **Efficiency** compares outputs with inputs and how economically inputs were transformed into outputs. It identifies waste, fraud and other inefficiencies. It weighs up the chosen approach against alternatives, examining not just whether the intervention has achieved its objectives but whether it has done so in a reasonably economic fashion. Could those resources have achieved more if used elsewhere or for a different function? It reviews both the manner in which resources were used and how well the implementation was organised and run. An efficiency evaluation will look for duplication of effort or gaps and whether there might have been a better way of achieving the same outcomes.

412. **Adaptability**.

a. **Responsiveness** examines how good an organisation is at recognising and responding to changes to the conflict dynamics; how well it can adjust its logic of change, plans and implementation.

The Purpose of Monitoring

413. Observable variables and indicators are monitored for four main purposes.

a. **Baseline Assessment**. Monitoring can be used to help populate or backfill a missing or incomplete baseline.

b. **Implementation Monitoring**. This tracks changes to lower level variables and objectives (input activity, output) identifying the extent to which the plan has been implemented. It may highlight areas were efforts need to be reinforced and provides basic data for an evaluation of efficiency (inside the shape).

c. **Progress Monitoring**. This tracks changes to higher level objectives (outcomes and impacts). It can indicate trends and provide insights into the validity of the intervention logic. It provides data for an evaluation of effectiveness and Impact (inside the shape).

d. **Assumption & Environment Monitoring**. This monitors changes to external factors that risk invalidating the logic of the plan. These may be specific to certain causal links or of a more general nature. Such changes should be identified quickly to allow plans to be adapted and enable a more responsive intervention (outside the shape).

Articulating the Key M&E Questions

414. Understanding the audience and the purpose of the evaluation allows key questions to be identified that will provide its focus. These should cover the three areas of "inside" and "outside" the shape and internally at ones own adaptability. The answers to these questions should provide the information that the client requires to make their decision and therefore ensure utility.

Achievable – Scope and Collaboration

Collaboration

415. Seeking ways to collaborate with others can provide a means of reducing the scope of the M&E task ensuring that it is achievable. It can also bring other benefits and risks.

416. **Avoiding Duplication and Exploiting Synergies**. It is important to explore what other M&E efforts are planned or underway within other organisations and at different levels within your own organisation and to seek opportunities for collaboration. Information from others may answer key questions allowing the scope the effort to be reduced. In doing this it is important to recognise that the logic tree of each organisation may have different reference points, what one considers an impact may be an outcome for another. Understanding how intervention logics relate to each other will suggest how different M&E efforts might link up.

417. **Collaborative Evaluations**. As well as the potential to reduce the scope of an evaluation, joint evaluations have a number of other significant benefits. Just as conducting a shared analysis of the problem can help bring different stakeholders together and generate a more coherent response so too can shared evaluations. The involvement of partners will bring fresh perspectives and is likely to improve the quality and coverage of the evaluation and, when the host nation is involved, can help build local capacity. Shared or dovetailed evaluations can reduce polling fatigue and may provide additional means of triangulating across different sources. Conducting evaluations with other stakeholders should always be considered.

418. **Multi-Level Evaluations**. Staff responsible for M&E planning should explore with colleagues at different levels of their own organisation opportunities to support each other and align their efforts. This may include aligning evaluation questions as well as the timing and means of data collection. It is particularly important for staff at higher levels to seek this dialogue, as much of the data collection will be undertaken by lower echelons that may have a different situational awareness and sense of priorities.

419. **Risks of Cooperation**. Decisions on collaborative evaluations need to balance benefits against the risks and costs of cooperating with other agencies. These include the risk of information leaking, training burden, differing priorities, prolonged discussion and the unintended consequences of involving some stakeholders but not others. There is also the

risk that by bringing stakeholders together a form a "group think" can develop whereby all fall into the trap of accepting the same assumptions. It may be wise to bring key stakeholders together as part of a joint evaluation while deliberately excluding some informed actors so as to compare and contrast perspectives. Typically NGO, Media or local civil society assessments may provide a useful comparator. Where assessments differ substantially alarms should sound and trigger work to explore the differences in perception.

420. **Risks of Shared Data**. Drawing on existing data also entails risk. It is essential to compare like with like to ensure that variations in data reflect variations in reality rather than variations in M&E methods. Problems may arise when external products are used as baselines against which in-house products are compared, or when data from different areas are compared. Care must be taken when aggregating data from different sources.

***IED Statistics in Afghanistan**. Until mid 2009, what warranted IED entries in the operational database varied between different ISAF units; some counted those that exploded, others unexploded but functional ones, or unexploded and dysfunctional ones, while yet others counted individual components found. The data was not comparable without going back to the original source reports and re-categorising them against common criteria.*

Participation

421. Participation in the evaluation needs to be decided early and should fall out of discussions on the level of collaboration. Two key decisions need to be made: who should be involved in the process and the extent of that participation. A question to be considered is the extent to which an evaluation is done *on* an organisation, environment or population or is done *with* them. At one level this highlights the balance between internal or external evaluation teams and between single or collaborative evaluations. At a deeper level it raises the question of the extent to which local officials, civil society and the population itself, those people who are the objects of the intervention, may be included in the evaluation.

422. **Internal or External?** There are pros and cons of using an internal, external or a mixed evaluation team. With little bias, no ownership of the plan and less stake in the result, external teams are particularly suitable for evaluations focusing on results and accountability. In contrast, internal teams have a good understanding of context, stakeholders, relationships, and the delivering organisation. This enables them to ensure that the evaluation is best tailored to support learning. However, they will be judging their own work and will have absorbed the implicit worldview of the organisation; biases may occur that undermine the integrity of the evaluation. Conversely recommendations from external evaluators that clash with the organisation's viewpoint can easily be dismissed and there is little chance for those evaluators to maintain pressure for change. Internal evaluators will however have developed a sense of ownership of the findings and are more likely to ensure they are followed through. A mix of internal and external evaluators may provide the most appropriate solution with external evaluators being used to expose and challenge implicit assumptions and perspectives.

423. **On or With?** Restricting involvement in the evaluation to those stakeholders actively conducting the intervention is a simpler and less risky approach which preserves a distinction between the intervention and the evaluation. It is however liable to reflect an external perspective and may misrepresent the actual dynamics of the situation. Including the

"objects" of the intervention, the local people, is likely to be a less controlled event that may struggle to provide quantitative data. But, their inclusion is likely to generate a sense of ownership and buy-in to the intervention that may substantially increase the quality of the data and the depth of the insights; it may also provide a springboard for any changes that are required, creating a strong foundation for subsequent activities. Conducting evaluations with the objects of the intervention makes the evaluation become part of the intervention contributing to the stabilisation effect.

424. **Level of Engagement**. As the membership and participation is developed it is equally important to define the extent of that participation and to clarify the roles of the participants. Will involvement simply extend to the sharing of information or will it be a form of consultation where views are sought but analysis and decision-making kept separate. Or will it be more of collaboration; an open and equal dialogue with participants involved in both analysis and decision-making. In either case the roles of participants need to be defined.

Prioritisation

425. M&E will always be constrained by resources or access to data and must therefore focus on areas where it can be most beneficial. Establishing clear priorities, drawing on existing work, collaborating with others and adjusting ambition to set realistic objectives is essential if the evaluation is to be achievable. Effective M&E is focused; optimising M&E is therefore more about clarity of purpose than scale of resources; large wide-ranging evaluations are not necessarily better. The purpose and key questions should be used to define the priorities for the assessment.

426. Ambition should be contained to what is realistic and necessary. While it might be useful to know how much each variable contributes to a particular outcome, allowing resources to be focussed where they can have most impact, this is likely to prove difficult in a stabilisation environment. Instead it may be more achievable to verify that it or that a group of variables are making a positive contribution to the outcome.

427. Priorities should be established to cover issues falling 'inside the shape', 'outside the shape' and within the organisation. Evaluators should be wary of ignoring any of these three areas as doing so could mean that critical causal links are not tested, vital changes in the environment not detected, or inappropriate processes not reformed. However while all three elements should be considered the balance between them may vary. It might not be necessary to assess all three at the same frequency or to the same depth.

428. **Inside the Shape**. Within the shape both causal links (connectors) and objectives (boxes) will be of interest. Tracking all these variables and assessing every link will usually be impossible and unnecessary. Effort should not be wasted testing and confirming causal pathways where confidence on their validity is high; the results will only reiterate the obvious and add no value. Evaluations should concentrate on high risk areas and test the empirical validity of:

a. Those causal links that will have the most significant bearing on overall objectives, and the ones which, if wrong, would severely impact on the intervention. This may include time critical objectives or those on which others are dependent.

b. Those causal pathways about which least is known.

c. Areas where little progress is made and uncertainty remains as to why.

429. **Outside the Shape**. M&E frameworks must also look beyond the "shape" and again concentrate on risk. While it may be appropriate to maintain a wide but shallow "horizon scanning" of external influences on conflict dynamics, efforts should focus on critical assumptions and high risk causal links in the theory of conflict that, if wrong, would undermine the change logic. For these high risk assumptions analysis should determine what might happen if they were wrong and specifically target those areas of the wider environment to provide early warning of potentially flawed logic.[7]

430. **Own Adaptability**. Recognising that the stabilisation environment is complex and difficult to predict it is important to assess one's own ability to sense changes and respond. Evaluation of the own organisation should therefore focus on one's information gathering and analysis processes as well as planning and decision-making processes. Within this it may be necessary to prioritise certain sectors. If the planning process suggests that sub-national governance is a key element of the intervention the assessment might focus on exploring how well existing structures and processes identify key information, work with key stakeholders and affect change across this sector.

431. **Monitoring Priorities**. While monitoring priorities will follow the priorities outlined above, the need to start collecting data as early as possible cannot be overemphasised. Priorities for data collection plans can be set in the following order:

a. Addressing gaps or irregularities in the baseline assessment. This will help inform the Theory of Conflict, improve the Change Logic, formulate realistic targets, and provide a starting point for subsequent comparisons.

b. Widen baseline assessment to cover impacts, allowing progress to be tracked.

c. Produce required data for foreseen evaluations.

d. Track inputs, activities and outputs to establish if plan is implemented. Note that the agency/units responsible for the delivery should be tasked to track these variables. This information may then be collated by those responsible for the monitoring.

Determining the Scope

432. Understanding the nature of other actors M&E efforts, the audience, purpose and priorities of your own efforts will help define the scope of the assessment identifying its breadth, depth and duration. Should it attempt to take a vertical slice through the logical tree focusing on a specific sector or cluster of objectives or should it take a broader systemic view? Alternatively should it take a horizontal slice across the logical tree focusing on a level of objectives such as outcomes rather than outputs? Should it focus on simply identifying the

[7] This is analogous to the identification of Named Areas of Interest (NAIs) within an Intelligence Preparation of the Battlespace (IPB) process.

immediate effects or should it look longer term and seek out the delayed impact of activities? In all cases it will be necessary to identify the appropriate balance between assessment inside the shape, outside the shape and assessment of oneself. The scope should also identify the key lines of inquiry to be followed, the granularity of the information required and the overall logic of the assessment.

433. **Partial or Systemic?** Conducting an assessment on a single sector or line of operation or within a smaller geographical area creates a more bounded problem and might be a more effective use of resources. However boundaries will be artificial constructs and issues within the sector are likely to be affected by external influences and drivers. In a more bounded evaluation it can be harder to identify these and consequently false lessons may be learnt about the level of contribution that the intervention has made and the accuracy or otherwise of the underlying hypotheses.

434. **Horizontal or Vertical?** Taking a vertical slice through a high risk or high priority area of the logical tree can focus efforts on confirming the validity of key causal links and the assumptions underlying them, proving the suitability of the approach. The risk is that other significant changes may not be identified. In addition higher level outcomes will take longer to appear and may not be discernible immediately. Taking a horizontal slice across the logical tree provides a comprehensive view across the intervention as a whole and is therefore more likely to identify unexpected developments but may mean that effort is wasted on low priority areas.

435. **Immediate or Long Term?** Typically the higher up the causal chain one focuses the longer it takes for effects to become apparent. Designers need to identify the time horizon at which they will be looking. To what extent are they going to focus on relatively short term, quick action/effect cycles with a view to making small and frequent course corrections and to what extent will they take a longer perspective making fewer but perhaps more significant course corrections. Focusing too much on the short term may make it hard to identify a subtle but significant change of direction over time while focusing too much on the longer term may mean that opportunities are missed.

436. **Granularity of Information**. It is important to identify the level of generalisation or disaggregation that will be required. Will it be necessary for example to break data down by region, ethnicity or religion? Greater levels of detail can provide a more nuanced understanding however if the data is to be statistically significant more samples will be required. There is a balance to be struck between the level of detail or granularity and the ability to collect the data. On the whole it is best to keep assessment approaches as complicated as required but as simple as possible. Keeping the client at the forefront of your mind will help insulate you from the siren voices of other secondary users and readers.

Timing, Duration & Frequency

437. The timing of the evaluation will depend on the purpose. If it is to inform a specific one-off decision the timing of that decision will drive the scope and approach of the assessment. In such cases, care must be taken to identify any seasonal factors that may affect the reliability of the results; weather patterns, agricultural cycles, social or religious events can all affect behaviour and distort the findings. If the assessment is part of a longer term

intervention management framework it will be possible to better take these into account and provide seasonally adjusted findings. Some activities are likely to lead to a rapid output and outcome. Others however may take longer to emerge. Those designing M&E frameworks should review initial findings as factors emerge subsequently; M&E is an iterative process of learning and so findings should rarely be considered complete or final – they are the best conclusions available at the time not necessarily the right ones!

Determining Lines of Inquiry

438. Clarity over the purpose, audience, key questions, and scope of M&E efforts will enable evaluators to focus efforts on broad themes or lines of inquiry. These should describe the main areas of examination and will guide the subsequent design.

Achievable – Designing the M&E Framework

439. The next step is to develop an M&E framework and plan that is achievable and will deliver the required information to the client.

Understanding and Prioritising the Intervention Logic

440. If M&E was considered during the planning phase the theory of conflict and the change logic should be clearly articulated and available. This is not always the case. Where it is missing evaluators may need to scour previous documentation to reconstruct the understanding at the time. If this is not feasible they should capture the logic as it stands at that moment. This will then form a baseline theory of conflict and change logic which can be used from then on as the basis for testing and refinement. Failure to do this will make meaningful evaluation impossible. Once the logic is understood evaluators will need to prioritise those areas as discussed in paragraphs 425 to 431. This will ensure that M&E efforts are tightly focused.

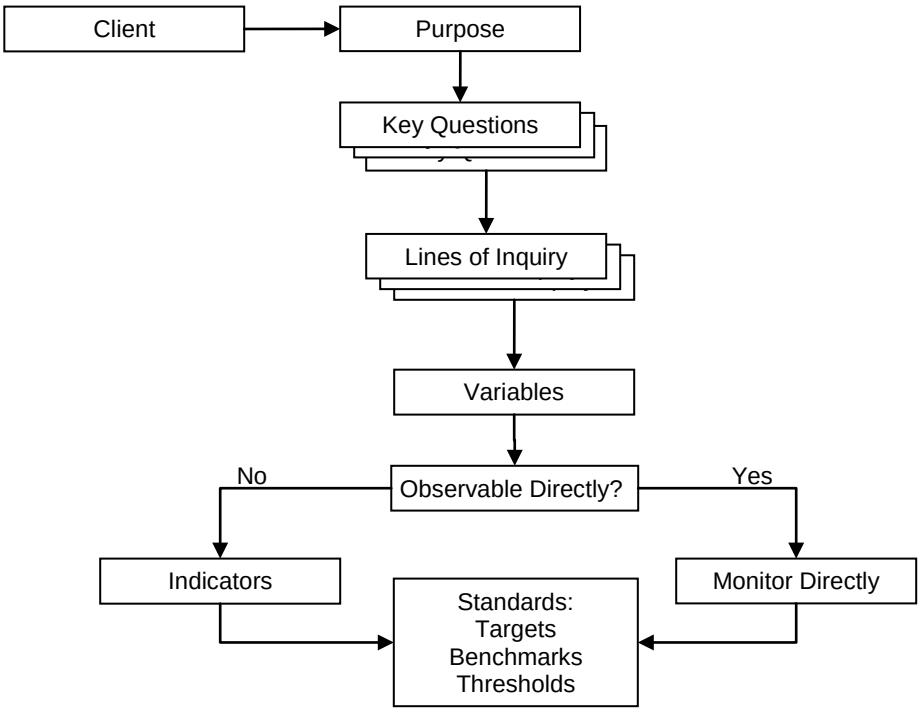
Identifying the Baseline

441. The baseline provides the starting point from which a comparison can be made and changes identified. Establishing the baseline is an important first step. This will require a review of existing data from a wide range of sources. It is not necessary to baseline everything but instead evaluators should focus their efforts on identifying as reliable a baseline as possible in those areas where the evaluation will focus. The nature of the stabilisation environment is such that there will be gaps, information will be out of date or inaccurate. Where these gaps are critical, effort may need to be dedicated to improving knowledge however it is often more practical to simply log what is readily available and use the M&E process to iteratively develop baseline data. This does mean that it will be harder to identify changes in the early stages of an intervention as the margin of error on data that does exist is likely to be substantial. In the early phases of an intervention it may therefore make sense to assume that the logic of the intervention is correct and to just monitor inputs and immediate outputs.

Identifying and Prioritising Variables of Interest

442. A variable is the issue or object that you wish to monitor the status of or measure change in. Variables can be deduced from the lines of inquiry and the key questions. Where these variables can be observed directly they become elements of the M&E framework.

Where they cannot be observed indicators must be identified that will suggest change in the variable of interest. Variables must be prioritised such that effort is focused on those that directly feed the key questions.



Establishing Indicators and Standards

443. Once the variables of interest have been deduced indicators need to be identified for those which cannot be observed or monitored directly. Standards may also be required for both indicators and directly observable variables. Identifying the right mix of indicators is a critical part of any evaluation. The following good practice should be kept in mind:

a. **Inclusion**. Involve key stakeholders in the identification and selection of indicators. It is particularly important to involve those who will collect the data and those whose expertise will be critical to the analysis of the data. Inclusion builds buy-in and is more likely to lead to quality data collection and support for the conclusions.

b. **Partnerships**. Establish partnerships with other key organisations and institutions as described in paragraph 416.

c. **Use Existing Data**. Existing data and reporting systems should be used wherever possible. If a state system exists it should be used. Where it lacks rigour or is open to abuse attempts should be made to work with it, triangulating from other sources. Bypassing it entirely may create parallel systems and undermine the credibility of the very institutions that you are trying to build. Working with it will help develop its capacity. Existing data can also provide a baseline and often provide an indication of trends over time. Be aware of the risks described in paragraphs 419 and 420.

444. Developing a set of indicators should follow an iterative process of identification, development and filtration. Evaluators are encouraged to adopt a “magpie” approach to developing sets of indicators with good ideas being adopted from anywhere and everywhere. These will need to be developed to make them relevant to the context and the purpose of the evaluation. At this stage evaluators should aim for a multiplicity of indicators which will then need to be filtered against the principles of utility, achievability, validity and conflict sensitivity. This process is discussed in more detail in Section 5.

Selecting Means of Verification

445. Selecting how the data will be collected, the means of verification (MoVs), should be addressed in parallel with the selection of indicators as this will be an iterative process. There is an inevitable tension between the desire for the ideal collection of indicators that provides sufficient balance and rigour with the practicalities of what can and cannot be measured. The aim should be to provide the most practical mix possible that achieves sufficient rigour while remaining achievable. Best practice suggests the use of multiple sources to triangulate perspectives and reduce the risk of bias. Possible approaches can be grouped into researching documents, asking people and observation. These are discussed in Section 5.

Selecting Means of Analysis

446. Once the data has been received it must be analysed; the crux of any evaluation is determining what the data means and what the implications are. Care must be taken to avoid making assertions that the data does not support, careful cross-referencing to the logic of the intervention and the logic of the evaluation can help. The analysis must seek out justifiable correlations between activities, outputs and outcomes constantly exploring what else might have caused or prevented the effect from being achieved; thinking in terms of contribution rather than attribution. The attitude should not be to demonstrate the accuracy of the intervention logic but to explore the logic, learning from the results and adapting it. Quantitative and qualitative analysis techniques are discussed in Section 5.

Forming the Evaluation Plan

447. Once variables, indicators, MoVs and the means of analysis have been identified it should be possible to draw the work together into an evaluation plan. This will lay out the audience and purpose of the evaluation along with the evaluation logic and its component parts. It should also provide the basis for tasking. An outline Evaluation Plans is attached at Annex **Error! Reference source not found.**.

Validity and Conflict Sensitivity – Testing the Framework

448. Having developed an achievable M&E framework and developed an evaluation plan that has utility (i.e. it meets the requirements of the client and satisfies the purpose of the evaluation) it is important to confirm that it is both valid and conflict sensitive. The framework should be tested to ensure that the data identified can indeed be accessed, that when analysed it provides responses that are useful in making decisions, that it is balanced and does not exhibit any unconscious bias and finally that it is conflict sensitive.

449. **Accessibility of Data**. Assumptions made over the availability of data or promises from organisations to supply data should be tested. It may turn out that organisations or individuals are less willing or able to share information in practice than theory or that the

information available is not of the quality required. Overcoming information sharing rules and managing mundane issues such as the interoperability of formats can all take longer than expected and early identification of these issues can save considerable time later.

450. **Utility of Response**. It is good practice to test the analysis process and presentation formats to ensure that it is capable of meeting the client's requirements and will enable them to make the required decisions. The framework should be populated with (fictitious) data that represents the possible extremes. This should be analysed and the conclusions presented to the decision-maker. Significant differences in data should lead to different conclusions, if they do not the framework is unable to provide the level of differentiation required and will need to be revisited.

451. **Balanced and Unbiased**. Inclusion of the right people in the development process should reduce the risk of bias nevertheless evaluators should expose their framework to a third party tasked with providing positive criticism. This may be an internally generated "red team"[8] or trusted outsiders, possibly embassy, DFID country office, UN agency or NGO staff that have not been involved in the process. Wherever possible a range of host nation people should also critique the process.

452. **Conflict Sensitive**. Finally the process should be examined to determine its potential impact on the conflict itself. It should not exacerbate the conflict drivers and if possible should reduce them. Conflict sensitive assessment is discussed in more detail in Section 5.

Conduct

Conducting the Analysis

453. It should be stressed that the analysis is not simply a question of summarising the data but is the process of drawing conclusions from it. While specialist statisticians and analysts have a vital role to play this should not be considered a "black art" and generalists must also be involved. The analysis should draw on the Evaluation Logic, the rationale that underpins the choice of indicators and standards to justify conclusions while recognising that this logic also should be tested. If results appear strange it may be that the indicator is not revealing what it was thought to reveal; it could be invalid. As the analysis progresses it is sensible to involve those who will have a role in staffing the conclusions or recommendations; their buy-in is likely to be essential if the recommendations are to be adopted and implemented with vigour. For large evaluations it may be wise to identify senior individuals as leads or champions for the areas that most affect them; these may be the people who will brief their areas of the evaluation to the client.

Reporting & Exploiting

454. The evaluation does not end with the presentation to the client and their decision. That decision and the recommendations that they endorse will need to be implemented. Evaluators should consider the exploitation plan for their analysis and ensure that time for re-

[8] Cross reference to the Red Team pamphlet

planning is built into the key organisations' "battle rhythms", that staff and resources are available and are synchronised with decision and planning cycles across these organisations.

455. In presenting the findings it is important to design the products so that they are appropriate for the audience and for the purpose. Understand how the client likes to receive information and the types of visualisation they prefer. Consider to what extent you will need to "sell" the conclusions and design the presentation appropriately. There is likely to be a tension between the desire to present the client with simple, clear messages and properly acknowledging the complexities and limitations of the findings that will need to be managed.

Changing the Framework

456. Over time the situation will develop and understanding of the environment will improve. As a result elements of the framework will no longer be relevant or will be proved invalid. The framework will therefore need to be adapted. Changing the framework should however be done with care as it may undermine the ability to track trends. There is a tension between resisting any changes so that trends and impacts can be discerned over time and the need to ensure that the evaluation remains both relevant and valid. Collecting data that is no longer a priority will limit the ability to collect data that is a priority and may also prove frustrating for those tasked with collecting it. This may undermine their willingness to engage with the process and hamper the ability to collect more significant data. The benefits of not changing the framework should be considered against the costs of continuing to collect and analyse data that is no longer a priority, may be of limited relevance and even misleading.

SECTION 5 – MONITORING & EVALUATION METHODS

Selecting Indicators & Standards

501. Selecting indicators and standards involves an iterative process of identification, development, filtration and testing.

502. **Identification**. Evaluators should adopt a "magpie" approach to developing sets of indicators and standards with good ideas being taken from anywhere and everywhere. It can be useful to develop an "indicator bank" of ideas. Evaluators should start with what has already been used by the organisation and by others across different sectors, levels and theatres. It can often be helpful to look at other professions beyond the immediate conflict area; public health, education, political science and psychology all seek to measure what are often intangible impacts and can provide useful ideas; possible sources are listed in Annex #. Asking key stakeholders what they see as being significant can also help focus in on the critical issues. Existing conflict and stakeholder analyses will also provide insights that can point to possible indicators.

503. **Development**. Development and filtration will overlap as the evaluation team discards possible indicators as their practicality or utility becomes better defined. Potential indicators should be made relevant to the context and to the purpose of the assessment. The constant refrain should be "can I gather this data, how reliable is it, will it help my primary user make the required decision?". As indicators are developed evaluators should also consider how the data will be collected. Although this can be refined later the discipline of asking the question should ensure that the identification phase remains grounded in what is practically achievable. A number of acronyms exist that can help develop and refine indicators and standards; these are shown in table # below.

Qualities of a Good Measure:	Components of an Indicator:	Qualities:	SMART:	QQT:
• Mapped to a desired objective or effect • Meaningful • Measurable • Culturally & Locally Relevant • Sensitive to Change • Time & Geographically Bounded • Comprehensive • Cost & Time Efficient	• What is to be measured • Unit of measurement to describe the change • Baseline status • Size of the intended change • Quality of the change • Target population • Timeframe	• Measurable • Reliable • Feasible • Useful	• Simple • Measurable • Achievable • Realistic • Timely	• Quantity • Quality • Time

504. **Filtration**. It is likely that the number of possible indicators will balloon at this stage and it is essential that they are filtered against the priorities and lines of inquiry so that a suitable balance is achieved. Constructing a matrix listing the indicators, identifying their importance for programme monitoring, the ease of obtaining data on the indicator and the cost of data collection may be helpful.

Indicator	Importance	Ease of Obtaining Data	Cost of Data Collection

The aim should be to achieve a mix of quantitative and qualitative indicators from a range of sources that together provide a sufficiently rigorous way of monitoring the status of each variable of interest. Care should be taken to ensure that data is not dependent on only one source and that the mix of indicators adequately covers the scope of the evaluation, picking up on objectives, causal links, the wider environment and our own structures and processes. In selecting indicators evaluators should consider what incentives these might create for subordinates, might the selected indicators encourage them to behave in a potentially perverse manner such that their activities impact on the indicators but not the variables?

505. **Testing**. Once the final basket of indicators and standards has been identified the evaluation framework should be tested. Will the data enable analysis and will the results of that analysis help the decision-maker make the required decision?

506. **Evaluation Logic**. In developing indicators a number of assumptions will be made about the linkages between indicators and variables that describe the logic as to why we think this indicator reveals the change in the variable that we are interested in. These linkages together form the Evaluation Logic and must be captured in exactly the same way as the Intervention Logic. Doing this will help those analysing the data once received, provide continuity on staff rotations and allow weak or flawed assumptions to be identified, tested and refined in subsequent iterations of the evaluation.

507. **Validity & Reliability**. Evaluations should be designed such that the results are both valid and reliable. Validity refers to whether the approach measures what it is supposed to measure; is it showing the changes that it was designed to illuminate or is it actually highlighting different dynamics. Reliability refers to the consistency of the results and consequently the dependability of the data and of the conclusions drawn from it. It asks whether the same results would have been obtained if the approach had been applied by different people from the same organisation and whether the same results would have been obtained if the process were repeated by the same people.

Gathering Data (Means of Verification)

508. **Triangulation**. When gathering data it is important to triangulate across different sources. Multiple theories, methods and data sources should be used to compare, contrast and substantiate information and opinions. Triangulation can identify and overcome any bias that may come from an individual source, observer, informant or interviewer. Several sources should be drawn from within one sample group (village, organisation, social group etc.) and should also be drawn from more than one group (several villages rather than one). Similarly several approaches should be used to collect data, for example survey data in itself may contain flaws however when matched with focus group interviews and direct observation these flaws may be overcome and the result have greater validity.

509. **Researching Documents**. Reviewing existing documents can shed light on the theory of conflict and the change logic that underpin an intervention and can provide insights into the thinking at the time and why certain actions were undertaken. These may well be different

from what actors state in the present as they will inevitably, and often unconsciously, have evolved their understanding and changed their perception. Reviewing other's evaluations and documentation can also highlight areas of agreement and disagreement that can support or challenge perceptions and prevent groupthink. Written sources should not be assumed to be "correct" or comprehensive; they are likely to contain errors and omissions and will reflect the bias and political context within which they were written. They are simply one form of evidence and should be compared and contrasted with evidence drawn from other means.

510. **Asking People**. Asking people can be as part of an interview or as a survey. There is a danger that respondents will simply tell the interviewer what they think they want to hear, more in depth interviews can help overcome this. In all cases careful design of the questions is vital. Questions can be open ended or closed; open-ended questions such as "How useful has the new road been to you?" allow people to report feelings, opinions and thoughts whereas closed questions such as "How often do you use the road?" provide easily collated and analysed data but can limit responses and conceal nuances. When asking people it is important to be clear what type of person you are asking as perceptions will vary.

An evaluation in Afghanistan conducted by the US think tank CSIS distinguished between three different types of "voice", beneficiaries, implementers and observers – where observers were individuals such as journalists. They found that observers tended to be more negative than either beneficiaries or implementers. The same evaluation also noted that responses from interviews were more negative than surveys. In interviews individuals tended to focus on the remaining challenges whereas the more structured survey forced them to focus on achievements as well.

a. **Interviews**. Interviews vary in terms of the number and type of people interviewed at once and the level of structure within the interview. In general the more structured an interview the more comparable the data however the more unstructured or open-ended the interview the more deeply interviewees feelings, understandings and perceptions can be ascertained.

b. **Surveys and Polls**. Surveys and polls comprise a focused set of targeted questions posed in a set sequence by a surveyor. They can provide a range of data rapidly from across a section of the population and are usually designed to provide statistically valid data.

511. **Observation**. Alternatively evaluators can record what they see and hear against an observation checklist. Observations may be of an ongoing process, activities, interactions or of the physical surroundings. Observation can help identify incorrect implementation of activities; it is also of use where security concerns make it difficult to interview people directly. It should be recognised that the presence of an observer, if known, will affect the situation and therefore what is being observed will no longer be "pure". The observer will also bring their own perspective and bias into the record.

512. **Advantages and Disadvantages**. The most common approaches are summarised below along with their advantages and disadvantages:

Researching Documents:

Type	Advantage	Disadvantage

Literature Search: Review of existing related documentation.	Economic and efficient way of obtaining information	Difficult to asses validity and reliability from secondary data
Case Studies: Investigate a contemporary event in its real life context.	Can provide a rich narrative and depth of understanding of a particular issue. Good for understanding processes and for formulating hypotheses to be tested later.	The choice of cases will influence the output. Large number of sources and perspectives will be required to ensure results are not distorted. It may not be easy to make generalisations about the conclusions without further evidence.
Diaries: Can be written, audio or video and provide a description of a personal experience. They can be structured so that individuals are focused on certain specific issues	A rich and varied sources that can provide important insights into perceptions and explain why decisions were made and activities conducted.	Time consuming to analyse and unlikely to provide statistically valid data. Little control of the content.

Asking People:

Key Informant Interviews: Interview with a key individual such as a village elder or local official. Can be structured, semi-structured or unstructured.	Flexible, in depth, easy to implement	Several individual interviews can be time consuming. Risk of bias, own agendas
Focus Group Interviews: Best conducted with 6-8 people who have something in common.	Reasonable in terms of cost and time, can be good for stimulating new ideas	Requires a skilled facilitator and a safe environment. There is a risk of one-sidedness and strong individuals may dominate the discussions; it may be difficult for individuals to dissent from the majority.
Group Interviews:	Low cost and efficient, direct contact with those affected	Susceptible to manipulation, less suitable for sensitive issues
Structured Interviews: Scope of the interview is pre-defined and a series of set questions are asked.	Easy to compare and contrast	Inflexible, important nuances may be lost.
Semi-Structured Interviews:	Flexible – richer and more varied insights	Less comparable and requires some skill in applying.
Unstructured Interviews: Interviewer is guided only by the evaluation objectives and the interview will flow to follow up areas of interest as they emerge.	Good for exploring opinions and uncovering unexpected issues.	Can be time consuming and difficult to compare and contrast.
Surveys:	Good for large numbers and for providing statistically valid data.	Time consuming; provides little in depth information about why things happen.
Questionnaires:	Easy to administer, can capture a wide population.	Need to be a statistically valid number. Question design is critical and poor questions can undermine the rigour of any conclusions. Requires a literate audience.

Observation:

Direct Measurement:	Precise, reliable and often requiring few resources	Requires access. Provides little qualitative information.
Direct Observation: Watching and taking notes, recording specific actions such as behaviours,	Can capture the experiences of minorities or women who may be unable to speak out due to cultural	Observers need to be able to access the location either directly or via surveillance assets. The results

attitudes, who was included or excluded, what interaction took place and between who. Observation can be supplemented by photos or videos.	norms.	depend on the observers' training, understanding and interpretation. Presence of the observer may also influence the situation. Does not explain why things happen.

Ways of Analysing & Using Data

513. Quantitative data analysis falls into the realms of statistical analysis and ideas such as statistical validity, distributions, significance and sample sizes all become relevant. A detailed discussion of statistical analysis tools is beyond the scope of this JDN. Any evaluation, particularly those with significant quantitative data, should be able to draw those with expertise in statistical analysis or run the risk of its findings being dismissed as invalid or lacking rigour. For qualitative data while some interview records or focus group discussions may provide stand-alone illustrations of important issues, it is important to analyse qualitative data more systematically. Such analysis might identify similarities between accounts, as well as directions, trends and tendencies that taken together may suggest more generalised conclusions. These may in turn form hypotheses for subsequent testing in the next iteration of interviews.

514. **Quantitative Analysis**. The strengths of quantitative approaches are that they can provide precise information that can easily be analysed to provide statistically valid conclusions. Data can easily be aggregated or disaggregated to provide a broad view of a population or of a small subsection of it and to enable comparisons to be made. However, while it may be precise it may not be measuring what was intended and this may not be immediately obvious; it can be precisely wrong. This is particularly the case within a stabilisation environment where problems of access, lack of host nation expertise and political drivers may make statistically valid data collection difficult. In such situations care must be taken that quantitative analysis approaches do not provide a veneer of respectability on what remains data of dubious validity. Similarly care must be taken in drawing conclusions from changes in quantitative data. What for example does a doubling of the number of open shops in a bazaar mean? If it doubles from 2 to 4 is this the same as it doubling from 8 to 16? And what do we mean by an open shop? Is a trader sat on the floor with a handful of wares for sale a shop? Care must also be taken in combining different factors and on applying weightings between them, what for example is the relative importance or significance of 2 km of rebuilt irrigation ditches as opposed to 2 km of new road? These decisions are context specific and subjective; analysing the implications of these requires a detailed, and probably local and tactical, level of knowledge. Quantitative approaches cannot explain the underlying causes of a situation, for this qualitative information is required.

515. **Qualitative Analysis**. Qualitative data and approaches are good at providing insights into attitudes, beliefs, motives and behaviours, particularly across small samples. They are vital for assessing progress and trends against qualitative objectives. Objectives such as the legitimacy of a local political leader depend on the perception of the local population and therefore need to be assessed through qualitative means. Qualitative data can often be gained rapidly and relatively cheaply however it is rarely statistically valid or representative of a wider population and therefore generalisations may not be valid. It is also more susceptible to influence by the perspective of those conducting the data collection, the interviewers, observers and informants. Although it can be appropriate to use quantitative analysis

techniques on qualitative data care should be taken not to take this to excess. Grouping responses, categorising them and reporting them in terms of ratios and percentages may mask important nuances and also prevent a more detailed and potentially more rewarding exploration of ideas, opinions and attitudes which might help answer the evaluation questions “why?” and “how?”.

Conflict Sensitive Assessment

516. The M&E effort is an intervention in its own right and has the potential to exacerbate tensions or reduce them. M&E must draw from the underlying theory of conflict and must consider the impact of the evaluation, its planning, conduct, reporting and exploitation, on the conflict drivers. The following questions should be considered:

a. What are the conflict drivers and how will this M&E affect them?

b. Who are the main conflict actors and how will this M&E affect them?

517. In answering these questions evaluators should consider the timing of the evaluation and its reporting. They should consider the choice of people to conduct the evaluation, including sub-contractors, and be on guard against any perceived bias or lack of balance. Similarly the choice of who is consulted or not consulted can have an impact; inclusion can either enhance or undermine the legitimacy of actors and therefore contribute to or undermine stability. It is important to consider the potential “winners and losers” in any evaluation and the impact this will have on the conflict. In the same way the choice of language or terminology can help or hinder; the way questions are phrased can reinforce or exacerbate tensions.

Existing Tools: Strengths and Weaknesses

518. A number of M&E tools exist and several of these are described in Annex #. Links are provided to more detailed descriptions.

ANNEX A: THEORIES OF CHANGE

The following are a selection of generic theories of change extracted from a number of sources[9].

- **Individual Change:** Peace comes through the transformative change in attitudes and behaviour of a critical mass of individuals.
- **Mass Attitudes**: If enough people's attitudes change then they will prefer that key actors seek peaceful solutions and will resist mobilisation to violence
- **Healthy Relationships & Connections:** Peace emerges from a process of breaking down polarisation and prejudice between groups.
- **Key Actor Attitudes**: If key actor attitudes change then they will seek peaceful solutions to conflicts.
- **Withdrawal of Resources for War:** Interrupting the resources and human capital required to sustain wars will cause it to stop.
- **Elite Means**: if the resources elites have to engage in organised violence are degraded or removed then they will be more likely to accept peace.
- **Reduction of Violence:** Peace will break out as the level of violence reduces.
- **Community Based Reconciliation**: If belligerent groups within a community are given the opportunity to interact, then they will better understand and appreciate one another and will prefer to resolve conflicts peacefully.
- **Root Causes / Justice:** Addressing the underlying grievances will take away the motivation for conflict.
- **Building Bridges**: If key actors from among belligerent groups are given the opportunity to interact, then they will better understand and appreciate one another, be better able to work with one another, and prefer to resolve conflicts peacefully.
- **Institutional Development:** Peace is secured through the development of stable and reliable social institutions that guarantee democracy, equity, justice and the fair allocation of resources.
- **Political**: If political institutions operated efficiently, impartially and in the interests of all, then the extent of core grievance would decline.
- **Political Elites:** Peace comes when it is in the interests of political and other leaders to take the necessary steps.
- **Elite Motivations**: If the incentives facing elites can be changed so that peace becomes more acceptable and violence less so then the elites will accept peace
- **Grassroots Mobilisation:** If enough people oppose the war their leaders will ultimately have to take notice of them.
- **Negotiated Settlement**: If we can establish space and mechanisms for negotiation between the leaders of the belligerent parties then they can be led gradually through a series of steps to cease violence and negotiate peace.
- **Economics:** Decisions are made on the balance of rewards/incentives and sanctions/punishments – shift this balance and peace will follow.
- **Economics**: If economic institutions produced reasonable livelihoods/quality of life for all, then the extent of core grievance would decline.
- **Public Attitudes:** War and violence are motivated by prejudice, misperceptions and intolerance, use the media etc. to change public attitudes and build greater tolerance.
- **Social Service Delivery**: If social services such as health care and education are delivered in an effective and responsive way for all, then the extent of core grievance would decline.
- **Security/Judicial**: If security and justice institutions protected everyone and enforced laws equitably, then the extent of core grievance would decline.
- **Culture of Peace**: If enough people's attitudes change then they will prefer that key actors seek peaceful solutions to conflicts and will resist mobilisation to adopt violence.
- **Transitional Justice**: If we create opportunities for members of war torn societies to come to terms with the conflict and their role in it and to heal the trauma they sustained, then the level of conflict related grievance will decline as will the likelihood of a return to war.

[9] Church & Rogers, Walker

ANNEX B: INDICATORS and EVALUATION TOOLS

Assessing a Country's General State – 'Outside the Shape'

B01. **The Conflict Assessment System Tool (CAST) and Failed states Index**. Originally published in 1996 and updated in 2007 by the Fund for Peace, the CAST manual provides an analytical model for early warning and risk assessment of weak and failing states based on twelve main and 38 subordinate conflict risk indicators, grouped under the headings Social, Economic, Political/Military. Although not tailored for a specific intervention they may be of some use when looking 'outside the shape'. By rating and aggregating the CAST indicators, the Fund for Peace produces an annual Failed State Index. Although outlining general trends for individual states, the utility of the Failed State Index to improve interventions is limited.[10]

a. The menu of CAST indicators is available at:

- http://www.fundforpeace.org/cast/pdf_downloads/castmanual2007.pdf

b. The Failed States Index is also available at:

- http://www.fundforpeace.org

B02. **The State Fragility Index (SFI)**. Developed by Monty G. Marshall and Jack Goldstone at the Center for Global Policy, George Mason University, the SFI focuses on the Polity of states and falls into the same category of tools. Results are published in the Global Report series, and are available at: The "State Fragility Index" was developed by Monty G. Marshall and Jack Goldstone at the Center for Global Policy, George Mason University, and has been an annual feature in the Global Report series, available at:

- http://www.systemicpeace.org/Global%20Report%202009.pdf

B03. **World Bank's Post-Conflict Performance Indicators (PCPI)**. Available at:

- http://siteresources.worldbank.org/INTCPR/1090479-1115613025365/20482305/Post-Conflict+Performance+Indicators,+2004-05.pdf

Selection of indicators – 'in and outside the shape'

B04. **Measuring Progress in Conflict Environments (MPICE)**. MPICE was developed in 2008 by the US Institute for Peace studies (USIP), the US Army Corps of Engineers and the US Army Peace Keeping and Stability Operations Institute. It offers a menu of more than 800 indicators and measures that can be used to capture change to variables that may be of interest in most interventions. The variables are organised in five themes: Political Moderation and Stable Governance, Safe and Secure Environment, Rule of Law, Sustainable Economy and Social well-being. As MPICE is based on the assumption that an intervention should diminish Drivers of Conflict and increase Institutional Performance, each theme contains variables in both these categories. The Menu of variables, indicators and measurements is available at:

10

- http://www.usip.org/files/resources/MPICE%20Aug%2008.pdf

Improving Theory of Conflict and Change Logic– bottom up approach

B05. **Tactical Conflict Assessment Framework (TCAF)**. Created by USAID, TCAF is a standardized diagnostic tool employed to gather information from local inhabitants to identify the causes of instability or conflict in tactical areas of operation. This information helps identify, prioritize, monitor, evaluate, and adjust civil-military programming targeted at diminishing the causes of instability or conflict. TCAF is based on asking locals four simple questions, each followed with 'Why?:

- Has the number of people in the village changed in the last year?
- What are the most important problems facing the village?
- Who do you believe can solve your problems?
- What should be done first to help the village?

TCAF has been subject of criticism, most notably for raising unrealistic expectations amongst local inhabitants. It does nevertheless, remain one of the few bottom up approaches that has been widely implemented. More info can be found at:

- http://www.usaid.gov/our_work/global_partnerships/ma/tcaf.html

B06. **Integrated Conflict Assessment Framework (ICAF)**. To Follow

B07. **State Department's Coordinator for Reconstruction and Stabilization (S/CRS)**. To Follow

Alternatives to the logical tree

B08. **Logical Frameworks and Results Frameworks**. Logical and Results frameworks represent the most commonly used ways of displaying the change logic and tend to identify indicators for all 'boxes' in the logical tree.

a. The World Bank (2004), by Jody Zall Kusek, Ray C. Rist, Ten Steps to a Results-Based Monitoring And Evaluation System. Available at:

- http://www.oecd.org/dataoecd/23/27/35281194.pdf

b. SIDA (2004), The Logical Framework Approach. A summary of the theory behind the LFA method. Available at:

- http://www.sida.se/shared/jsp/download.jsp?f=SIDA1489en_web.pdf&a=2379

c. DFID (2009), 'Guidance on using the revised Logical Framework', How to note, A DFID practice paper. Available at:

- http://mande.co.uk/blog/wp-content/uploads/2009/06/logical-framework.pdf

Conflict Analysis Tools

B09. A list of conflict analysis tools can be found at:

- http://www.gsdrc.org/go/conflict/chapter-1-understanding-violent-conflict/conflict-analysis-framework-and-tools

ANNEX C: OUTLINE EVALUATION PLAN

Purpose

C01. Describe the purpose of the evaluation.

Audience

C02. Describe the primary and subsidiary audiences

Decision and Key Questions

C03. Describe the decision(s) that the evaluation is seeking to inform and outline the key questions that the evaluation will answer in order to inform that decision.

Scope

C04. Describe the scope of the evaluation including the balance between “inside”, “Outside” the shape and “Self”.

Priorities

C05. State the priorities for the evaluation.

Collaboration and Participation

C06. Highlight linkages with other evaluations and define the level of collaboration with other organisations and across levels of your own organisation.

C07. Define who will be involved in the evaluation and the level of participation they will have.

Timing

C08. Define the timing, duration and frequency of the M&E effort.

Lines of Inquiry

C09. List the lines of inquiry that will be pursued.

M&E Framework

C10. Outline the framework of indicators, standards, MoVs and MoA linking them to the lines of inquiry and key questions.

Indicator	Standard	Means of Verification			Means of Analysis
		Method	Data Source	Collection Task	
Key Question: …					
Line of Inquiry: …					

Data Collection Plan

C11. Summarise the data collection plan and allocate tasks for collection and collation of data.

Analysis Plan

C12. Summarise the analysis plan and allocate tasks for the fusing and analysis of data.

Evaluation Logic

C13. Describe the causal links between selected indicators, the variables of interest, the lines of inquiry and key questions.

M&E Management

Information Management and Exploitation

C14. Describe how the data will be managed, and exploited, including information security issues allocating tasks and responsibilities.

Presentation

C15. Describe how the evaluation will present its findings concentrating on the strategic communications, media and influence issues.

Conflict Sensitivity

C16. Highlight key conflict sensitivity concerns and approaches.

Risk Management

C17. Describe the major risks, how they will be monitored and mitigated.

GLOSSARY

- **Activity**: Actions taken or work performed through which inputs, such as funds, technical assistance and other types of resources are mobilised to produce specific outputs.
- **Assumption**: Hypotheses about factors or risks which could affect the progress or success of an intervention
- **Attribution**: The ascription of a causal link between observed (or expected to be observed) changes and a specific intervention. A statement outlining the proportion of observed change which can really be attributed to the evaluated intervention. An attribution analysis normally involves a control group or a counterfactual scenario.
- **Baseline**: An analysis describing the situation prior to an intervention, against which progress can be assessed or comparisons made.
- **Benchmark**: Reference point or standard against which performance or achievements can be assessed.
- **Causality**: A definitive statement that one event is caused by a second event.
- **Collaborative Evaluation**: An evaluation conducted with one or more external organisations or agencies. The level of collaboration may vary.
- **Contribution**: A statement that one event has helped bring about a second event. It is likely to be one of a number of causes not all of which have been identified. A contribution analysis may also rank the evaluated intervention among the various causes explaining the observed change.
- **Correlation**: The tendency for two variables or events to change together. The stronger the correlations the closer the two variables co-vary. A strong correlation does not, however, necessarily mean that a cause and effect relationship exists. A strong correlation between atmospheric CO2 levels and crime levels does not of course mean that atmospheric CO2 causes crime.
- **Effectiveness**: The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.
- **Efficiency**: A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.
- **Evaluation**:
- **Ex-Ante Evaluation**: An evaluation that is performed before implementation of a development intervention.
- **Ex-Post Evaluation**: Evaluation of a development intervention after it has been completed.
- **Formative Evaluation**: Evaluation intended to improve performance, most often conducted during the implementation phase of projects or programs.
- **Generalisation**: The extent to which one can ascribe conclusions from one set events in one context to another set of events in a different context. Will for example conclusions from an evaluation of actions taken in one region of a country apply if similar actions were undertaken in another region of the country.
- **Goal**: The higher-order objective to which an intervention is intended to contribute.
- **Impact**: Positive and negative, primary and secondary long-term effects produced by an intervention, directly or indirectly, intended or unintended. Reduction in crime; increase in literacy.
- **Indicator**: Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of an actor.
- **Input**: The financial, human and material resources used for the intervention.
- **Meta-Evaluation**: The term is used for evaluations designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators.
- **Metrics** are a system of measures that allow one to assess the status of issues of interest. The term is often used as a collective noun for the baskets of indicators, means of verification and means of analysis that make up an evaluation approach.

- **Monitoring**: A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.
- **Outcome**: The short-term and medium-term effects of an intervention's outputs, i.e. the exposure to or use of outputs. Police officers pursuing criminals or manning checkpoints; pupils taught in school.
- **Output**: The products, capital goods and services which result from an intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes. E.g. police officers trained; schools built.
- **Performance**. The degree to which an intervention or a partner operates according to specific criteria/standards/ guidelines or achieves results in accordance with stated goals or plans.
- **Process-Evaluation**: An evaluation of the internal dynamics of implementing organizations, their policy instruments, their service delivery mechanisms, their management practices, and the linkages among these.
- **Program-Evaluation**: Evaluation of a set of interventions, marshalled to attain specific global, regional, country, or sector development objectives.
- **Project-Evaluation**: Evaluation of an individual development intervention designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader program.
- **Purpose**: The publicly stated objectives of the program or project.
- **Redundancy**: The inclusion of multiple data sources to mitigate the risk of missing or corrupt information and thus ensure the validity of the evaluation.
- **Results**: The output, outcome or impact (intended or unintended, positive and/or negative) of an intervention.
- **Summative Evaluation**: A study conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced. Summative evaluation is intended to provide information about the worth of the program.
- **Triangulation**: The use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment.
- **Validity**: The extent to which the data collection strategies and instruments measure what they purport to measure.

BIBLIOGRAPHY

Recommended Reading: Top Documents

1. Church, C. and Rogers, M. (2006), Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs. Available at:
   - www.sfcg.org/Documents/manualintro.pdf
2. OECD/DAC (2008), "Guidance on Evaluating Conflict Prevention and Peacebuilding Activities – Working Draft for the Application Period", Available at:
   - http://www.oecd.org/secure/pdfDocument/0,2834,en_21571361_34047972_39774574_1_1_1_1,00.pdf
3. SIDA (2004), Stefan Molund and Göran Schill, Looking Back, Moving Forward, Sida Evaluation Manual. Available at:
   - www.sida.se/PageFiles/3736/SIDA3753en_Looking_back.pdf

Literature:

M&E Manuals/ Guidance

4. Asian Development Bank 2006, Impact Evaluation- Methodological and Operational Issues. Available at:
   - http://www.adb.org/Documents/Handbooks/Impact-Analysis/Impact-Analysis.pdf
5. DFID, Guidance on Evaluation and Review for DFID Staff, July 2005
   - http://www.dfid.gov.uk/Documents/publications/evaluation/guidance-evaluation.pdf
6. DFID, M&E A Guide for DFID Contracted Research programmes, May 2006
   - http://www.dfid.gov.uk/Documents/publications/me-guide-contracted-research.pdf
7. NONIE (2009), Impact Evaluations and Development. Available at:
   - http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf
8. OECD (2007), Encouraging Effective Evaluation of Conflict prevention and Peacebuilding Activities. Towards DAC Guidelines.
9. OECD DAC (2002), "Glossary of Key Terms in Evaluation and Results Based Management", DAC Network on Development Evaluation, OECD, Paris.
10. OECD/DAC (2006) Evaluating Development Co-operation. Summary of key norms and standards.
11. OECD INCAF Team on Peacebuilding, Statebuilding, Security, "Draft Justice and Security Monitoring and Evaluation Toolkit" Nov 2009
12. UNFPA The Programme Managers Planning, Monitoring & Evaluation Toolkit
    - http://www.unfpa.org/monitoring/toolkit.htm Aug 2004
13. USAID Direct Observation

- http://pdf.usaid.gov/pdf_docs/PNABY208.pdf

14. USAID Focus Group Interviews
    - http://pdf.usaid.gov/pdf_docs/PNABY233.pdf
15. USAID Key Informant Interviews
    - http://pdf.usaid.gov/pdf_docs/PNABS541.pdf
16. USAID Rapid Appraisal Methods
    - http://pdf.usaid.gov/pdf_docs/PNABS541.pdf
17. USAID, Monitoring & Evaluation in Post-Conflict Settings (Draft), March 2006
18. World Bank (2006), Conducting Quality Impact Evaluations under Budget, Time and Data Constraints. Available at:
    - http://lnweb18.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/757A5CC0BAE22558852571770059D89C/$file/conduct_qual_impact.pdf
19. World Bank (2006a), Impact Evaluation- The experience of the independent evaluation group of the World Bank. Available at:
    - http://lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/35BC420995BF58F8852571E00068C6BD/$file/impact_evaluation.pdf

M&E (Case) studies

20. ABCA, “Measuring Effectiveness of Military Contribution to Security Sector Reform” Report by Capability Group ACT Project Team 09ACT002 Oct 2009
21. Ryan, S. and Hiscock, D. “Evaluating for Security and Justice: Challenges and opportunities for improved monitoring and evaluation of security system reform programmes” Saferworld, Oct 2009

Methodology

22. Babbie, Earl (2001), The Practice of Social Research, 9th edition, Wadsworth London.
23. Bernard, Russel (2000), Social Research Methods. Qualitative and Quantitative Approaches, Sage London.
24. Burns, Robert (2000), Introduction to Research Methods, 4th edition, Sage London.
25. CSIS, Breaking Point: Measuring Progress in Afghanistan March 2007
26. Pedhauzur, Elazar and Schmelkin, Liora (1991), Measurement, Design, and Analysis: An Integrated Approach, Lawrence Erlbaum Associates London.
27. McNamara, Carter (2002), A basic Guide to Programme Evaluation. Available at:
    - http://www.tgci.com/magazine/A%20Basic%20Guide%20to%20Program%20Evaluation.pdf
28. Sholz, Rolland and Tietje, Olaf (2002), Embedded Case Study Methods. Integrating Quantitative and Qualitative Knowledge, Sage London.

29. Stern. Elliot (ed) (2005), Evaluation Research Methods, Volume 1, Sage London.
30. Summer, Andy and Tribe, Michael (2008), International Development Studies. Theories and Methods in Research and Practice, Sage London.

Doctrine

31. DCDC, JDP 3-00
32. DCDC, JDP 3-40 “Security and Stabilisation: The Military Contribution”, November 2009
33. DCDC, JDP 5-00 “Campaign Planning” Second Edition, December 2008
34. LWC, Doctrine Note 09/07 “Measuring the Effectiveness of Operations”
35. US Joint Chiefs of Staff Joint Publication 3-07 “Stability Operations” Final Coordination Draft, April 2010
36. US Joint Chiefs of Staff Joint Publication 5-0 “Joint Operation Planning” Revision First Draft, January 2010

General

37. Anderson, M.B. (1999a), Do No Harm: How Aid Can Support Peace or War, Lynne Rienner, London.
38. Mavrotas, George and McGillivray, Mark (eds.) (2009), Development Aid. A Fresh look, United Nations University, Palgrave macmillan Basingstoke.

Web sites

39. OECD/DAC Evaluation of Development Programmes:
    - http://www.oecd.org/department/0,3355,en_2649_34435_1_1_1_1_1,00.html
    - http://vision2020.info.tt/cms/index.php?option=com_content&task=view&id=76&Itemid=85
    - http://www.unfpa.org/monitoring/toolkit/5data.pdf
    - http://www.usaid.gov/policy/ads/200/203.pdf
    - http://www.unfpa.org/monitoring/toolkit.htm
40. Monitoring and Evaluation News
    - http://mande.co.uk/